ISSN 2279-9362

The Best in the Class

Ainoa Aparicio Fenoll

No. 644
April 2021

Carlo Alberto Notebooks

www.carloalberto.org/research/working-papers

The Best in the Class

March 16, 2021

Abstract

I estimate the effect of being the *best in the class* in primary school on performance in secondary school. I implement a novel methodology that does not require experimental data. My methodology exploits that some students are the *best in the class* because better students in the same school are assigned to other classes. If students were randomly assigned, the probability of being the *best in the class* would be a well-known function of students ranking in the school and the number of classes. I use this exogenous probability as an instrument for actually being the *best in the class*. I find a positive impact of being the *best in the class* on future performance: being the best in primary school increases test scores by 0.13 standard deviations in secondary school.

1 Introduction

Top-performing students are likely to be tomorrow's leaders. Jeff Bezos, founder and chief executive officer of Amazon.com and richest man in the world, graduated valedictorian of his high school and summa cum laude from Princeton University. Bill Gates, founder of Microsoft and second richest man in the world, was a National Merit Scholar and scored 1590 out of 1600 on the SAT (Forbes, 2019). Also according to Forbes (2000), most top CEOs excelled in education: 24% of top CEOs in Europe have PhD degrees, while the proportion of PhD graduates among top CEOs in China is 33%. However, little research is devoted to excellent students. In the US, the National Association for Gifted Students regrets the absence of a uniform federal policy for "gifted services". This lack of regulation results in a variety of State policies which go from "Accommodation in the regular classroom" (where gifted students become the best in the class) to "Full-time grouping with students of similar abilities" (where, with one exception, excellent students are not the best in the class).[1] In this paper, I study the consequences of being the best student in the class for future academic performance.

Previous literature shows that, on average, higher positions in the school-cohort ranking lead to better future academic performance. However, this average effect may not apply to the top performer in the class: first, being *the best* is more salient. Students may know who the best student in their class is, but may ignore the identity of the fifth or the nineteenth student. Visibility may imply social approval which boosts self-confidence and therefore future performance (Ferkany, 2008). It may also imply more attention and better treatment by teachers and peers. However, it may also imply higher expectations about students' future performance which may harm those students with low capacity to cope with pressure (Cotton, Dollard, and De Jonge, 2002). Second, the very best student may be demotivated from lack of competition, worsening future performance. Third, the very best students do not have better peers who may influence them positively, which puts them at a disadvantage in comparison to similar students who have better peers.[2]

[1] See https://www.nagc.org/ for more detailed information.

[2] The peer effects literature shows the positive influence of high ability students on their peers: Sacerdote [2001], Whitmore [2005], Carrell, Fullerton, and West [2009], Black, Devereux, and Salvanes [2013], and Booij, Leuven, and Oosterbeek [2017].

Finally, best students may be more likely to be bully victims, which harms future performance (Brown and Taylor, 2008).

I propose a novel methodology to estimate the impact of being the best in the class on future performance that does not require experimental data. I exploit the allocation of students to classes within a school. In practice, this allocation may not be random. For this reason, I use the theoretical probability of being the best in the class under random assignment of students to classes within a school as an instrument for being the best in the class. My identification strategy relies on two facts: first, a student becomes the best in the class only if all better students in the same school-cohort are allocated to other classes. Second, if the allocation of students to classes within a school was completely random, the probability of being the best in the class would be a deterministic function of the student's position in the school-cohort ranking and the number of classes in the school-cohort. For example, the second student in a school-cohort with two classes only becomes the best in her class if the first student is allocated to another class. This occurs with a probability of one half. If the second student in the school-cohort attends a school with three classes, this probability goes up to two thirds. Finally, the third student in the school-cohort has a probability of being the best in the class in a school with two classes equal to the probability that both the first and the second student in the school-cohort are in the other class (0.5^2=0.25).

My objective is to compare the future performance of individuals with the same ability where one is the best in the class and the other one is the second best. To make these individuals comparable, I use information on students' performance in a standardized test. Previous literature takes into account individual's ability by including a polynomial of test scores in their regressions. Instead, I take into account individuals' ability by including test score fixed effects and show that my results are consistent when I use the highest rank polynomial used by previous literature.[3]

An additional challenge for the estimation of best in the class effects is that the assignment of students to classes may not be random. I account for selection into classes as

[3]Current test scores also include contemporaneous effects of being the best in the class and hence my estimation of the effect of being the best in the class on future performance is net of those contemporaneous effects.

well as peers and teacher characteristics by including class fixed effects in my regressions. The fixed effects absorb all mean differences between classes but they do not account for differences in the variance and higher moments (skewness, kurtosis, etc.) of the ability distribution between classes. For instance, principals may group individuals with similar abilities or the worse students with the best ones and still get the same mean ability. To counteract the influence of these higher moments, my instrument is based on information from the entire school-cohort which is unaffected by selection of students into classes.

I use data on standardized tests administered to all students in Italy. These tests cover two subjects (mathematics and reading), are designed by an agency of the Italian Government (the National Institute for the Evaluation of the School System - INVALSI), and are mandatory for all students. Students are tested in second, fifth, eighth, and tenth grades of compulsory schooling. In my main analysis, I use information on mathematics test scores for fifth and eighth grades, which correspond to the last grades of primary school and secondary school.

Best in the class effects could be consistently estimated by OLS regressions of future performance on the best in the class dummy including class fixed effects if the assignment of students to classes within a school was random. I test whether my data is consistent with random allocation by studying whether being the best student in the school-cohort makes it more likely to be in class with other highly-ranked individuals. I also use the random allocation test proposed in Guryan, Kroft, and Notowidigdo [2009], which explores the correlation between test scores within a class and hence, it is not specific to the study of ranking effects. Both tests indicate that there is a small positive correlation between students' performances within classes in my data. On one hand, the positive sign of the correlation indicates that allocation of students to classes within a school is not random and therefore I cannot interpret OLS estimates of future performance on best in the class as causal. This justifies the use of my instrument. On the other hand, the small magnitude of the correlation implies that there are a significant fraction of students who were randomly assigned to classes. As my instrument provides an estimate of the best in the class effect for randomly-assigned students (the so-called compliers), my IV estimates are arguably representative of the population.

Results show that being the best in the class (as opposed to being the second) significantly increases students' future performance. The magnitude of the estimated effects is high. Being the best in the class in fifth grade increases test scores by 0.13 standard deviations in eighth grade. Results are robust to the use of data from other grades and reading rather than mathematics test scores. Estimates cannot be explained by students' absences or measurement error in students' performance. The effect of being second rather than the third in the class on future performance is also positive but smaller in magnitude.

1.1 Related Literature

I focus on the best students in the class. The economic literature that studies talented students is scarce. Some exceptions are Griffith and Rask [2007] and Horstschräer [2012], who study talented children's school choice and Figlio and Lucas [2004], who analyze the impact of high grading standards on high ability students.

This paper relates to the literature on the impact of relative position in the school ranking on future educational outcomes. Recent papers on this argument include Murphy and Weinhardt [2020], Elsner and Isphording [2017], and Denning, Murphy, and Weinhardt [2018]. Murphy and Weinhardt [2020] find that being ranked highly during primary school has large positive effects on secondary school achievement in the UK, with the impact of rank being more important for boys than girls. To identify ranking effects, they account for test scores in primary school using polynomials up to cubic and include school-subject-cohort fixed effects in their regressions.

Elsner and Isphording [2017] and Denning, Murphy, and Weinhardt [2018] focus on long run ranking effects. In the US context, Elsner and Isphording find that if two students with the same ability have a different rank in their respective cohort, the higher-ranked student is significantly more likely to finish high school, attend college, and complete a 4-year college degree. Denning, Murphy, and Weinhardt find that a student's rank in third grade negatively impacts grade retention while it positively affects test scores, high school graduation, college enrollment, and earnings up to 19 years later in the US. In contrast to the three papers mentioned above, I estimate the effect of being the best rather

than the average effect over all ranking positions.[4] Moreover, I focus on the best in the class rather than the best in the school.

The closest papers to mine focus on the class ranking rather than the school ranking and exploit a random allocation of students. Cicala, Fryer, and Spenkuch [2017] find that, in the context of 61 Kenyan primary schools, increasing a student's class rank by fifty percentiles boosts test scores at the end-line by about 0.2 standard deviations. They make use of random allocation of students to classes within the same school and assume that ability is well accounted for using a quadratic polynomial of test scores. Bertoni and Nisticò [2018] exploit a similar experiment implemented in the University of Amsterdam where first year students in Economics were randomly allocated to tutorial groups. They show that students with higher ordinal ability rank within groups have better academic outcomes. In their setup, moving from the bottom to the top of the within-group ability distribution increases the number of credits achieved by about half of a standard deviation. Elsner, Isphording, and Zölitz [2018] exploit data with repeated random assignment of students to teaching sections and find that a higher rank increases performance and the probability of choosing related follow-up courses and majors. In my paper, I focus on the best student rather than the average effect of ranking positions and my methodology does not require experimental data.

Given that *the best in the class* is a very salient position, my paper closely relates to the literature on ranking concerns. There is evidence suggesting that students care about their achievement rank even in the absence of specific rank incentives (Tran and Zeckhauser, 2012, and Azmat and Iriberri, 2010). Rank concerns have been studied also in various fields outside of education, for example, in the study of well-being at work and job satisfaction (Brown, Gardner, Oswald, and Qian, 2008, and Card, Mas, Moretti, and Saez, 2012), of performance in sports tournaments (Genakos and Pagliero, 2012), and of labor market productivity (Vidal and Nossol, 2011), among others. Tincani [2017] points at ranking concerns as one of the mechanisms behind the heterogeneity of peer effects.

The remainder of this paper is organized as follows. I present the data and institu-

[4] Murphy and Weinhardt [2020] have separate parameters for being top or bottom in the school-cohort and find discontinuously large estimates for being at the extremes.

tional background in Section 2. In Section 3, I describe my methodology and in Section 4, I present my results. Section 5 discusses several extensions and robustness checks. I conclude in Section 6.

2 Data and Institutional Framework

Education is compulsory in Italy from age 6 to 16. The education system is divided into primary school (five years), secondary school (three years), and high school (five years). Admission to Italian primary schools is based on a point system in which distance from home to school, having attended a kindergarten under the same school administration, and number of siblings (especially if they already attend the same school) increase the likelihood of admission. The weight given to each of these factors changes across municipalities. I provide more details about the institutional framework in Appendix A.1.

I use standardized test score data from the National Institute for the Evaluation of the School System (INVALSI) which covers the universe of Italian students. Students take standardized tests in the second quarter of the second and fifth year of primary school, then three years later in the third year of secondary school and finally two years later in the second year of high school. INVALSI provides data from academic years 2009–10 to 2016–17.

INVALSI tests present two crucial features for this analysis: first, all students in Italy take the same test. This allows me to rank individuals in the same school and to use standardized scores as a measure of ability. Second, individual identifiers are available for the academic years 2013–14 to 2016–17. I use these identifiers to link individuals in two consecutive tests taken three years apart. This is crucial for my identification strategy because I can study the impact of being the best in the class in a given grade on performance three years later: performance in eighth grade (secondary school) of the best in the class in fifth grade (primary school).

From second to fifth grade most students remain in the same school, with the same classmates and teachers. In contrast, from fifth to eighth grades all students change school, teachers and at least part of their classmates. For the identification of ranking

effects, it is crucial to rely on a pre-determined test score. Murphy and Weinhardt [2020] for example use primary school (KS2) test scores to rank students in secondary school. Elsner and Isphording [2017] rank students on a measure of crystallized intelligence which they argue to be fixed before primary school. I hence focus on the impact of being best in school in primary school (fifth grade) on performance in secondary school (eighth grade).

The INVALSI data contains test scores from two subjects (reading and mathematics) and indicates the number of correct answers. I focus on mathematics tests instead of reading because Italian proficiency is likely to be affected by migrant status or the use of a regional language at home (14% of students declare to speak a language other than Italian at home). Still, I check the consistency of my results with those obtained using reading test scores in Section 5. I standardize test scores by subject, academic year, and grade to have zero mean and unit variance (as in Angrist, Battistin, and Vuri, 2017). The data set also includes students' characteristics (among them: gender, migrant status, and whether they attended daycare and/or kindergarten) and parental characteristics (among them: migrant status, level of education, and occupation).

I make a series of exclusions to arrive at the sample that I use for my analysis. I start from individuals who attended fifth grade in the academic years 2012–13 to 2013–14 because those are the only cohorts for which I have information on test scores in eighth grade. From this sample, I select individuals who are in the first eighteen positions of the school ranking. I choose that threshold because for those individuals the probability of being the best in their class is at least 1%. I show that results remain arguably unchanged when I use individuals in the first ten, fifteen, and twenty-five positions instead. Finally, I exclude students in schools with only one class as my instrument is not exogenous for them.

The resulting data set includes $249,309$ students. As part of my supplementary analysis, I also estimate the impact of being the best in second grade on fifth grade and the impact of being the best in eighth grade on tenth grade for which there are $234,502$ and $147,520$ students, respectively.[5] The average fifth grade student answers correctly 75%

[5]Education is compulsory up to age 16. Thus, some non-randomly selected students drop out by tenth grade. This attrition could bias my estimation of the impact of best in the class in eighth grade on performance in tenth grade. I nevertheless show these results in Section 5. They are consistent with those

of the questions in the mathematics tests. The corresponding percentages for second and eighth graders are 76% and 81%, respectively.

Table 1 presents average key characteristics of students and their parents. I describe separately the samples used in the estimations of the effect of being the best in the class in second grade on fifth grade performance (first two columns), being the best in the class in fifth grade on eighth grade performance (third and fourth columns) and being eighth in the class on tenth grade performance (last two columns). I first comment on the samples of fifth to eighth graders which are used in my main analysis and then highlight the differences with respect to the sample of second to fifth graders and eighth to tenth graders.

The average test score in the sample of fifth to eighth graders moves from 0.8 standard deviations in fifth grade to 0.6 in eighth grade. The average student in my sample has a theoretical probability of being the best in the class slightly below 0.15. On average students are in schools with 3 classes and 18 students per class.[6] As my sample is composed of students in ranking positions $1 - 18$, the average student in my sample is number 9 in the school-cohort ranking. There are slightly more males than females. Although the incidence of foreign-born is relatively low (3%), around 8% of students have an immigrant father and around 10% of mothers are immigrants. In the sample of second to fifth graders, descriptive statistics are extremely similar to those in the main estimation. Only the reduction in test scores three years later is sharper. In the sample of eighth to tenth graders, average test scores are much higher (1.14 in eighth grade which go down to 0.77 in tenth grade). Also, the theoretical probability of being the best in the class is much higher (0.23). The latter is a consequence of the larger number of classes in the average school (4). Demographic characteristics are comparable across the three samples.

Table 11 in Appendix A.2 provides further information for these groups; it describes daycare and kindergarten attendance and the education and labor market status of parents. The proportion of students who attended daycare was 34% among fifth graders. As much as 86% of students attended kindergarten. Regarding the education level of par-

obtained when using students in compulsory education.

[6]There are 62% of students in schools with two classes, 21% in schools with three classes, 11% with four classes, 5% with five classes, and the remaining 1% in schools with six or more classes.

ents, 43% of mothers have a high-school degree. The proportion of university graduated mothers is 21%. Fathers are slightly less educated: 39% of them have a high-school degree and around 18% have a university degree. The proportion of homemakers among mothers is relatively high (33%). Although the proportions of white collar workers are the same for mothers and fathers ($43 - 44\%$), the proportions of self-employed and blue collar workers are low for mothers (9% and 11% in each category). These proportions are much more relevant for fathers (one fourth of fathers are self-employed and another fourth are blue collar). These characteristics are in line with those observed for the sample of second to fifth and the sample of eighth to tenth graders except for the proportion of students who attended daycare which decreases with the grade (moves from 39% to 34% and 27%) and for the proportion of students who attended kindergarten which is higher (89%) for children in the highest grades.

Table 1: Descriptive Statistics. Grades 2, 5 and 8.

	Grade 2		**Grade 5**		**Grade 8**	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Test score in t	0.769	0.639	0.778	0.663	1.137	0.623
Test score in $t + \tau$	0.481	0.840	0.606	0.879	0.771	0.926
Instrument	0.148	0.265	0.148	0.265	0.232	0.294
# classes	2.647	1.019	2.65	1.026	4.178	2.339
# students in class	19.135	3.761	18.458	3.772	19.683	3.926
Student ranking in grade	9.444	5.184	9.449	5.186	9.247	5.189
Male	0.531	0.499	0.534	0.499	0.541	0.498
Immigrant child	0.019	0.138	0.027	0.163	0.031	0.173
Immigrant father	0.085	0.28	0.079	0.27	0.069	0.253
Immigrant mother	0.108	0.31	0.099	0.299	0.088	0.284

Notes: This table presents averages and standard deviations (left and right column, respectively) for each sample used in the estimations. The number of observations is $234,502$ in grade 2, $249,309$ in grade 5 and $147,783$ in grade 8. $t + \tau$ refers to the next grade for which an INVALSI test is available.

3 Methodology

My first objective is to address whether students are assigned to classes depending on their ranking in the school-cohort.[7] If allocation was completely random, I could estimate the impact of best in the class using an OLS regression of future test scores on best in the class controlling for class dummies. Otherwise, my instrument becomes useful to estimate the impact of best in the class on future performance. The random allocation test also speaks about the properties of my IV estimation. If class assignment is close to a random allocation, the correlation of best in the class with the exogenous probability of being the best in the class is high and hence my instrument is strong. My instrument provides estimates of the best in the class effect for students who are randomly assigned to classes. Therefore, if class assignment is close to random, the group of randomly-assigned students is wider and my estimates are more representative of the population.

I explore the randomness of class assignment by estimating the probabilities that the first student is in class with the second, third, fourth, etc. student in the school-cohort ranking. In practice, I regress a dummy for best in the school-cohort on dummies for one of the classmates being in ranking positions two to eighteen as follows:

$$\begin{aligned} \textit{Best in cohort}_{i,t} &= \beta_0 + \beta_1 D(\textit{2nd in class})_{i,t} + \beta_2 D(\textit{3rd in class})_{i,t} + ... \\ ... &+ \beta_{17} D(\textit{18th in class})_{i,t} + \beta_{18} \textit{School-cohort}_{i,t} + \beta_{19} \textit{Class-size}_{i,t} + \beta_{20} D(t) + u_{i,t} \end{aligned} \tag{1}$$

where *Best in cohort* is a dummy for student i being the best in the school-cohort at time t and t takes values 2012-13 and 2013-14. $D(\textit{n-th in class})$ are dummies equal to one if the student in ranking position number n in the school-cohort ranking is in class with student i. I also include dummies for *School-cohort*, indicators for number of students in the class, *Class-size*, and year dummies, $D(t)$. Standard errors u are clustered by school. I run the regression separately for second, fifth, and eighth grades.

[7] I test for randomness of class assignment with respect to the observed ranking which is different from the innate ability ranking or the pre-school ranking. For instance, principals may have initially assigned students randomly with respect to their pre-school ranking but peer effects, teacher traits, and school characteristics may generate a non random allocation in fifth grade. Alternatively, principals may have sorted students across classes according to their ability but later changes in the ranking may result in a more random allocation.

Under random assignment, we expect the β coefficients to equal minus the proportion of students that are first ranked in their class, but not in the school-cohort. The reason is that the probability that a student is in class with the best student in the school-cohort is negatively affected by the fact that one of the slots in the class of the best student is occupied by her. Guryan, Kroft, and Notowidigdo [2009] propose a correction to account for this negative bias in the context of a standard random assignment test. The standard test is a regression of individual test scores on the average test score of the class excluding the individual. As in the previous test, this automatically generates a negative bias. Guryan, Kroft, and Notowidigdo [2009] correct this bias by including the average test score of all potential class members excluding the student as a control in the regression. The resulting equation is as follows:

$$TS_{i,t} = \gamma_0 + \gamma_1 Mean\ TS\ in\ class_{-i,t} + \gamma_2 Mean\ TS\ in\ school_{-i,t} + \gamma_3 School\text{-}cohort_{i,t} + ... \\ ... + \gamma_4 Class\text{-}size_{i,t} + \gamma_5 D(t) + v_{i,t} \tag{2}$$

where TS is the mathematics test score at the individual level while *Mean TS in class* and *Mean TS in school* are the average test scores of all students excluding individual i in the class and school-cohort, respectively. Under random assignment, γ_1 equals zero.

If the coefficients arising from the estimation of equations (1) and (2) are equal to their theoretical values under random assignment, I can estimate the impact of being the best in the class on future outcomes using OLS. Otherwise, I need to use an IV. In this latter case, if the coefficient estimates are close to their theoretical values, class assignment is close to random, and as a consequence, my instrument is strong and my identification strategy provides representative estimates of the effect of best in the class on future performance.

After this preliminary check, I move to the estimation of the impact of being the best in the class on future performance. I regress the test score in secondary school on a dummy

for being the best in the class in primary school as follows:

$$TS_{i,t+3} = \alpha_0 + \alpha_1 Best_{i,t} + \alpha_2 Third_{i,t} + \alpha_3 Fourth_{i,t} ... + \alpha_{17} Eighteenth_{i,t} + ... \\ ... + \alpha_{18} X_{i,t} + \alpha_{19} D(TS)_{i,t} + \alpha_{20} D(Class)_t + \varepsilon_{i,t} \tag{3}$$

where *Best* is an indicator of best student in the class. *Third*, *Fourth*,..., and *Eighteenth* are indicators for positions 3 to 18 in the class ranking. These dummies leave being second in class as the reference category and hence allow me to define α_1 as the effect of being best in class as opposed to second in class. X are the individual characteristics described in tables 1 and 11, $D(TS)$ are dummies for test scores (rounded up to the first digit), and $D(Class)$ are class fixed-effects.

Including students' test scores at time t accounts for students' ability. Contemporaneous test scores also control for contemporaneous effects of being the best in the class on test scores. Cicala, Fryer, and Spenkuch [2017] account for ability including the baseline score and its square, Murphy and Weinhardt [2020] include polynomials up to the cubic term, while Elsner and Isphording [2017] include a fourth order polynomial of test scores in their regressions. In my main analysis, I use the most flexible specification by first rounding the test score data up to the first digit and then including (rounded) test score fixed effects in my regressions. Finally, the vector of class fixed effects is necessary to account for average selection into classes, peer effects, teacher quality, and any other average unobservable characteristic which is common to all students in a class.

In the context of the previous regression, some concerns on the exogeneity of best in the class may arise. The influence of selection into classes may operate through higher order moments rather than the average. The best student in the class may have a very different experience in a class with average students and in a class with half excellent and half mediocre students. For this reason, I propose an instrument that provides consistent estimates even if class allocation is not fully random. The instrument is also useful to attenuate measurement error or in case specific types of students move classes. It also helps if expectations about future test scores influence the effort students make to become the best in the class.

3.1 The Instrument

My instrument is the theoretical probability P of being the best in the class under random assignment of students to classes within a school. To construct P, I take two variables as given: students' position in the school-cohort ranking and the number of classes in the school-cohort. Therefore, P is the theoretical probability that all better students in the same school-cohort are allocated to other classes. In practice, this is a deterministic function of position in the school-cohort ranking and number of classes in the school-cohort such that:

- If the student is the *second* in the school-cohort & there are *two* classes in the school, then $P = 1/2$.
- If the student is the *second* in the school-cohort & there are *three* classes in the school, then $P = 2/3$.
- If the student is the *third* in the school-cohort & there are *two* classes in the school, then $P = 1/2 * 1/2 = 1/4$
- If the student is the *third* in the school-cohort & there are *three* classes in the school, then $P = 2/3 * 2/3 = 4/9$

The general formula for this theoretical probability is:

$$P = \left(\frac{\#classes - 1}{\#classes}\right)^{(CR-1)} \tag{4}$$

where $\#classes$ is the number of classes in the school-cohort and CR is the position of the student in the school-cohort ranking.[8]

Figure 1 shows the theoretical and empirical probabilities of being the best in the class. The empirical probabilities are the proportion of best in the class students among those

[8] In computing the theoretical probability of being the best in the class I omit that this probability depends on the number of slots in each class. Results remain invariant when I use an alternative instrument that takes class size into account. I construct this alternative instrument under the condition that all classes within a school-cohort are equally sized. Thus, I avoid that the correlation between students' ability and class size affects my instrument. Such correlation is different from zero if, for instance, principals assign poor performing students to small classes.

with the same theoretical probability. The 45-degree line illustrates the reference case in which theoretical and empirical probabilities are the same, i.e., under random assignment. The actual probabilities of being the best in the class are similar but typically smaller than the corresponding theoretical values under random assignment. This can be explained if in some cases, students of similar abilities are assigned to the same class or if there are peer effects.

The instrument is a function of the school-cohort ranking and the number of classes in the school. The instrument would be invalid only if there was an omitted factor correlated with the specific functional form in Equation (4) after controlling for class fixed effects.

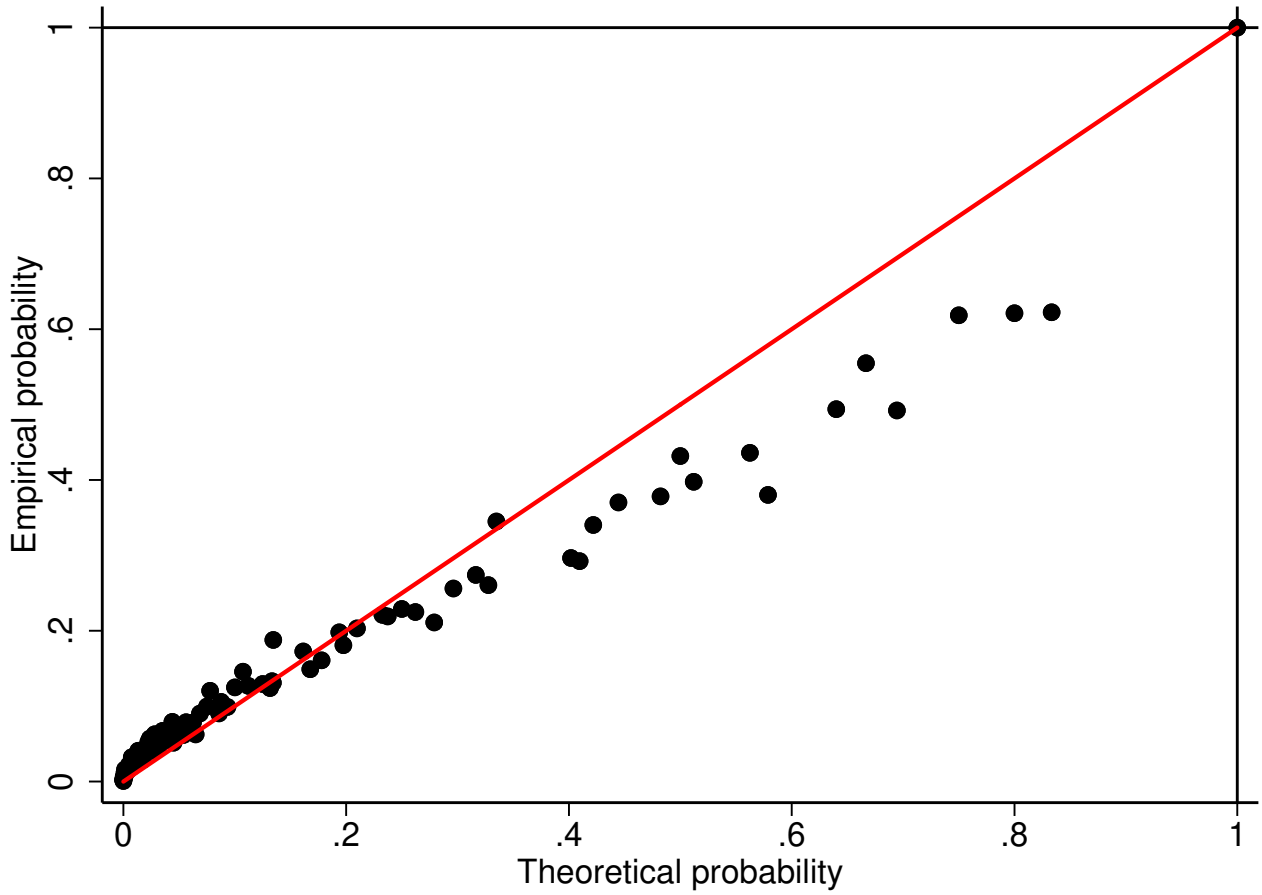
4 Results

I first present the results of testing the random assignment of students to classes within a school. I then show the naïve OLS estimates of test scores on best in the class which constitute a reference for the causal estimates. Finally, I describe the results of the set of regressions associated with the causal estimate of best in the class on future performance.

Table 2 shows the results of the estimation of Equation (1). For assignment to be random the estimated coefficient should equal minus the proportion of students that are first ranked in their class, but not in the school-cohort. This number equals -0.034. Hence, under random assignment, the coefficients associated with the second, third,..., eighteenth ranked student being in the same class as the best student in the school-cohort should be -0.034. The estimated coefficients are higher than the target values and become more negative as students are ranked lower in the school-cohort ranking. This indicates that there is some positive assortative matching across classes within the same school, especially at the very top of the school ranking. This positive assortative matching may arise because principals assign students of similar ability to the same class or because of peer effects.

I next present the results of the test for exogeneity of assignment of individuals to different groups proposed by Guryan, Kroft, and Notowidigdo [2009]. This test incorporates a correction of the negative bias induced by the presence of the individual herself in the

Figure 1: Empirical versus theoretical probabilities of being the best in the class



Notes: Data is from INVALSI test for the years 2012-13 and 2013-14. Each dot represents the proportion of best of the class students for each value of the theoretical probability of being the best in the class.

Table 2: Random Assignment Test

	Grade 2	Grade 5	Grade 8
Second in class	0.006	0.007	-0.001
	(0.001)***	(0.001)***	(0.001)
Third in class	0.001	-0.001	-0.005
	(0.001)	(0.001)	(0.001)***
Fourth in class	0.001	-0.0003	-0.006
	(0.001)	(0.001)	(0.001)***
Fifth in class	-0.0004	-0.003	-0.006
	(0.001)	(0.001)***	(0.001)***
Sixth in class	-0.001	-0.002	-0.010
	(0.001)	(0.001)**	(0.001)***
Seventh in class	-0.005	-0.004	-0.009
	(0.001)***	(0.001)***	(0.001)***
Eighth in class	-0.004	-0.004	-0.007
	(0.001)***	(0.001)***	(0.001)***
Ninth in class	-0.004	-0.0054	-0.012
	(0.001)***	(0.001)***	(0.001)***
Tenth in class	-0.004	-0.007	-0.011
	(0.001)***	(0.001)***	(0.001)***
Eleventh in class	-0.007	-0.006	-0.013
	(0.001)***	(0.001)***	(0.001)***
Twelveth in class	-0.008	-0.006	-0.012
	(0.001)***	(0.001)***	(0.001)***
Thirteenth in class	-0.008	-0.008	-0.016
	(0.001)***	(0.001)***	(0.001)***
Fourteenth in class	-0.007	-0.009	-0.015
	(0.001)***	(0.001)***	(0.001)***
Fifteenth in class	-0.008	-0.011	-0.014
	(0.001)***	(0.001)***	(0.001)***
Sixteenth in class	-0.008	-0.009	-0.011
	(0.001)***	(0.001)***	(0.001)***
Seventeenth in class	-0.007	-0.011	-0.013
	(0.001)***	(0.001)***	(0.001)***
Eighteenth in class	-0.009	-0.011	-0.015
	(0.001)***	(0.001)***	(0.001)***
Obs.	234,616	249,579	147,894
R^2	0.002	0.003	0.006

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14. The dependent variable is a dummy for being the best in the school-cohort. All regressions include school-cohort fixed effects, class size dummies, and year indicators. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

analyzed group. The results of this test suggest that there is a small positive correlation among students' performance within classes. This confirms that I need to use my instrument to obtain consistent estimates of the effect of being the best in the class on future performance.

Table 3: Random Assignment Test with Negative Bias Correction

	Grade 2	Grade 5	Grade 8
Mean TS class	0.013	0.021	0.015
	(0.002)***	(0.002)***	(0.007)**
Mean TS grade	-31.472	-29.478	-10.525
	(0.171)***	(0.189)***	(0.135)***
Obs.	234,616	249,579	148,894
R^2	0.897	0.89	0.856

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14. The dependent variable is the standardized mathematics test score. All regressions include school-cohort fixed effects, class size dummies, and year indicators. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

I then move to the estimation of the causal effect of being the best in the class. As a reference point, I estimate the OLS regression of test scores in secondary school on being the best in the class in primary school controlling for a fourth order polynomial of test scores at t, indicators of other positions in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators (see column 1 of Table 4). I first add class fixed effects (column 2) and then substitute the fourth order polynomial of test scores by test score dummies (column 3) to arrive to the specification in Equation (3). We find that the association between best in the class and future test scores is positive and small in all regressions. Controlling for class fixed effects reduces the coefficient associated with best in the class from 0.10 to 0.07. Substituting the fourth order polynomial in test scores by test score fixed effects leaves the coefficient almost unaltered. This suggests that unobserved average characteristics of the class may be biasing the OLS coefficient upward.

To address the potential endogeneity of best in the class in the OLS regressions above, I use the instrument defined in Section 3.1. The first stage estimations displayed in Table 5 show that the instrument is strong in all three specifications. The magnitude of the coeffi-

Table 4: Test Scores in Secondary School on Best in the Class in Primary School

	(1)	(2)	(3)
Best in class	0.097	0.068	0.069
	(0.005)***	(0.005)***	(0.005)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.245	0.254	0.254

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14 together with test scores for years 2015-16 and 2016-17. The dependent variable is the mathematics test score in secondary school. All regressions include dummies for positions 3-18 in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

cients indicates that more than 92% of students in my sample are compliers, i.e., for them being the best in the class depends on the exogenous probability that better students are assigned to other classes.

Table 5: First Stage. Best in the Class in Primary School on Theoretical Probability

	(1)	(2)	(3)
Theoretical probability	0.718	0.92	0.922
	(0.004)***	(0.003)***	(0.003)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.553	0.627	0.627

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14. The dependent variable is a dummy for being the best in the class. All regressions include dummies for positions 3-18 in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Next, I estimate the impact of the theoretical probability of being the best in the class in primary school on test scores in secondary school. I use a reduced form specification in which I substitute the dummy for being the best in the class by the theoretical probability of being the best in the class under random assignment in Equation (3). The results of such exercise are displayed in Table 6. The effect of the theoretical probability of being the best in the class on future test scores is positive, significant and consistent across

specifications. The magnitude of the estimated causal effect shows that a change in the theoretical probability of being the best in the class from zero to one increases test scores three years later by 0.12 standard deviations.

Table 6: Reduced Form. The Impact of Theoretical Probability of Being Best in the Class in Primary School on Test Scores in Secondary School

	(1)	(2)	(3)
Theoretical probability	0.118	0.117	0.121
	(0.01)***	(0.01)***	(0.01)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.244	0.254	0.254

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14 together with test scores for years 2015-16 and 2016-17. The dependent variable is the mathematics test score in secondary school. All regressions include dummies for positions 3-18 in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

To gain a sense of the magnitude of the impact of being the best in the class on future test scores I instrument being the best in the class by the theoretical probability of being the best in the class. Results in Table 7 show that individuals who become the best in the class because school mates who are better than them are assigned to a different class improve their test scores three years later by 0.13 standard deviations on average. These results are robust to the use of raw test scores computed as the sum of correct answers rather than standardized test scores. Results become even stronger when I exclude the best student in the school-cohort who does not contribute to the identification. Removing students with test score equal to that of another classmate leaves the estimated effect of best in the class arguably unchanged (there are 7% of test score ties in the data.).

5 Robustness Checks and Extensions

In this section, I first study the robustness of my main results. In particular, I check whether students who are best and second in the class are comparable in terms of predetermined characteristics after controlling for all the variables in Equation (3). Next, I

Table 7: The Impact of Best in the Class in Primary School on Test Scores in Secondary School. IV

	(1)	(2)	(3)
Best in class	0.165	0.127	0.132
	(0.014)***	(0.011)***	(0.011)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.244	0.253	0.253

Notes: Data is from INVALSI test for the years 2012-13 and 2013-14 together with INVALSI test scores for years 2015-16 and 2016-17. The dependent variable is the mathematics test score in secondary school. All regressions include dummies for positions 3-18 in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Best in the class is instrumented using the theoretical probability of being the best in the class. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

explore whether my results are affected by students' absenteeism. I then study whether my results are driven by measurement error in students' performance.

I also extend my analysis in several ways: I study the impacts of being the best in the class in second grade on performance in fifth grade and being the best in the class in eighth grade on performance in tenth grade, I replicate my analysis using reading test scores rather than mathematics test scores, and I explore the impacts of being second and worst in the class on future performance.

Next, I perform some heterogeneity analysis. In particular, I study how the estimated effect changes by gender, by class size, and by school quality. I also explore whether the best in the class effect changes when I modify my sample to include students at different positions of the school-cohort ranking. Finally, I study how the estimated effect changes with the distance between the best and the second student in the class.

5.1 Robustness Checks

My estimation strategy assumes that being best rather than second in the class is as if randomly assigned in the context of my regressions. If this is the case, there should not be any effect of best in the class on pre-determined characteristics. Panel A of Table 8 shows the result of using each of the predetermined characteristics in tables 1 and 11 as alternative dependent variables in Equation (3). All estimated effects are highly insignif-

icant except for the effect of being the best in the class on the male dummy which gives a negative significant estimate. I argue that this is not a concern since males tend to get higher mathematics grades and hence it operates against the positive effect of best in class on future performance. Moreover, I control for a male dummy in all regressions.

In my analysis, the main regressor of interest is a dummy equal to one if the student is the best in the class. In practice, the best in the class is the student with the highest test score in the class. One potential concern is that the variable best in the class is subject to measurement error if the actual best student in the class is absent the day of the test or has a "bad day". To understand the consequences of this source of measurement error in terms of coefficient estimates, I randomly select five percent of students in my original estimation sample, exclude them from the sample, and replicate the main estimation. The results of that exercise are shown in Panel B of Table 8. They show that the absence of randomly chosen individuals from the sample leaves the coefficient almost unchanged.

Students' performance on the INVALSI test may not mirror their school performance. INVALSI data contains information on the mathematics score given by the teacher in the first quarter of the academic year. This score is a discrete number from one to ten. It is an intermediate score which does not appear in the final student record. Hence, it is also an imperfect measure as teachers may use it to affect students' effort. I use the score obtained in the first quarter as an alternative measure of students' school performance. 82% of best in the class students according to INVALSI tests are also best in the class according to the score obtained in the first quarter. In Panel C of Table 8, I restrict my sample to those students who are best in the class according to both scores and results become even stronger.

5.2 Extensions

As mentioned above, INVALSI data also allows me to study the impact of best in the class in second grade on performance in fifth grade (both in primary school) and the impact of best in the class in eighth grade (third grade of secondary school) on performance in tenth grade (second grade of high-school). The transition from second to fifth grade takes place

Table 8: Robustness Checks

Panel A: Effect of Best in the Class on Pre-determined Characteristics. IV

	male	immigrant	immigrant father	immigrant mother
	(1)	(2)	(3)	(4)
first-class	-.044	-.0004	0.005	-.004
	(0.008)***	(0.002)	(0.004)	(0.004)
Obs.	247,471	247,471	247,471	247,471
R^2	0.012	0.165	0.325	0.339

	mother high school	father high school	mother university	father university
	(1)	(2)	(3)	(4)
first-class	0.011	0.002	0.005	0.005
	(0.007)	(0.007)	(0.006)	(0.005)
Obs.	247471	247471	247471	247471
R^2	0.112	0.121	0.225	0.228

	mother employed	father employed	mother unemployed	father unemployed
	(1)	(2)	(3)	(4)
first-class	0.007	0.01	-.005	-.002
	(0.007)	(0.007)	(0.003)	(0.003)
Obs.	247471	247471	247471	247471
R^2	0.198	0.244	0.013	0.03

Panel B: Excluding Randomly Selected Students. IV

	(1)	(2)	(3)
Best in class	0.166	0.127	0.131
	(0.015)***	(0.012)***	(0.012)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	237015	237015	237015
R^2	0.244	0.253	0.253

Panel C: Scores given by the Teacher in the First Quarter. IV

	(1)	(2)	(3)
Best in class	0.233	0.195	0.199
	(0.014)***	(0.012)***	(0.012)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	243288	243288	243288
R^2	0.25	0.259	0.259

Notes: Data is from INVALSI test for the years 2013-14 and 2014-15 together with test scores for years 2015-16 and 2016-17. The dependent variable in panels B and C is the mathematics test score in secondary school. All regressions include dummies for positions 3-18 in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Best in the class is instrumented using the theoretical probability of being the best in the class. Standard errors are clustered at the school level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

in the same school so it is more difficult to consider the class ranking in second grade as pre-determined with respect to fifth grade performance. Still, I present the corresponding results in Panel A of Table 9. Similarly to the transition from primary to secondary school, the transition from eighth to tenth grade implies a change in school, from secondary to high school. Unfortunately, the estimation of best in the class effects on performance in tenth grade is affected by attrition as some students in that grade become sixteen and can therefore legally drop out from school. As those students are unlikely to be at top positions of the school-cohort ranking, I show the results of such estimation in Panel B of Table 9. Both sets of results confirm the positive impact of being the best in the class on future performance. The estimated effects are stronger as compared to those in the main analysis and indicate that being the best in the class reduces future test scores by 0.15 standard deviations.

Throughout my analysis, I focus on mathematics instead of reading test scores because the former are less affected by Italian language proficiency which varies significantly across Italian regions depending on the incidence of regional languages and migration. Still, I replicate my analysis using reading test scores to define the ranking of students and to measure their performance. Panel C of Table 9 shows the results of this exercise which are stronger than those obtained using mathematics test scores. The magnitude of the estimated effects indicates that being the best in the class in primary school increases reading test scores in secondary school by 0.17 standard deviations.

Are ranking effects linear? If they are, the impact of being best in the class as opposed to the second best should be equivalent to the impact of being second as opposed to third in the class. I estimate the impact of being the second in the class on future performance using the theoretical probabilities of being second in the class as an instrument (see Panel D of Table 9). The general formula for these theoretical probabilities is:

$$P = \left(\frac{1}{\#classes}\right) * \left(\frac{\#classes - 1}{\#classes}\right)^{(CR-2)} \tag{5}$$

for all school-cohort ranking positions other than the first. For the best student in the school-cohort P equals zero. All specifications show that being second in the class in pri-

mary school has a positive and significant effect on performance in secondary school. In terms of magnitude, the impact of being second in the class equals 0.08 standard deviations. Therefore, the impact of being best rather than second in the class is higher than the impact of being second rather than third in the class by 65%. This constitutes evidence against the linearity of ranking effects. My findings are in line with the non-linear relationships found in experiments that emphasise the importance of being ranked first or last (Gill, Kissová, Lee, and Prowse, 2019 and Kuziemko, Buell, Reich, and Norton, 2014).

The methodology employed to estimate the impact of the best of the class on future performance can be used to estimate the impact of being the worst in the class. In this case, the instrument is based on the probability that individuals who are worse than a given student are in classes different from him or her. The estimated coefficient is close to zero and insignificant and hence, I could not find any effect of being worst in the class as opposed to being second worst. Results are available upon request.

5.3 Heterogeneity

Similarly to the average ranking effect (Murphy and Weinhardt, 2013), the effect of being the best in the class may differ across genders. I explore this possibility by interacting the *best in the class* variable with a male dummy. Results, presented in Panel A of Table 10, are consistent with Murphy and Weinhardt [2013] in that males are more affected by ranking effects. Males who are best in the class increase their future performance by 0.08 standard deviations more than females in the first position of the class ranking. I also explore whether the estimated positive effect differs by class size and school quality as measured by the average score in the school. I find that the best in the class effect decreases with class size and increases with school quality.

In my main analysis, I restrict my sample to students who are in positions from one to eighteen in the school-cohort ranking. Those students have probabilities above one percent of being the best in the class. Moreover, eighteen is also the average number of students in a class. In this section, I explore how my results change when I modify this sample selection criterion. Panel B of Table 10 contains the results of using three different

Table 9: Extensions

Panel A: The Impact of Best in the Class in Second Grade on Test Scores in Fifth Grade. IV

	(1)	(2)	(3)
Best in class	0.233	0.158	0.154
	(0.013)***	(0.01)***	(0.01)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	234502	234502	234502
R^2	0.177	0.2	0.201

Panel B: The Impact of Best in the Class in Secondary School on Test Scores in High School. IV

	(1)	(2)	(3)
Best in class	0.219	0.145	0.153
	(0.019)***	(0.015)***	(0.015)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	147783	147783	147783
R^2	0.218	0.233	0.233

Panel C: The Impact of Best in the Class in Primary School on Reading Test Scores in Secondary School. IV

	(1)	(2)	(3)
Best in class	0.215	0.163	0.168
	(0.011)***	(0.009)***	(0.009)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249195	249195	249195
R^2	0.211	0.22	0.22

Panel D: The Impact of Second in the Class in Primary School on Test Scores in Secondary School. IV

	(1)	(2)	(3)
Second in class	0.029	0.073	0.079
	(0.021)	(0.018)***	(0.018)***
Best in class	0.14	0.131	0.137
	(0.013)***	(0.013)***	(0.013)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.245	0.254	0.253

Notes: Data is from INVALSI test for the years 2013-14 and 2014-15 together with test scores for years 2015-16 and 2016-17. All regressions include dummies for other positions in the class ranking (excluding the reference category), the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Best in the class is instrumented using the theoretical probability of being the best in the class. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

samples including students who are in ranking positions up to ten, fifteen, and twenty-five. As a result of changing the sample, the number of observations becomes 135,020, 205,868, and 247,471, respectively. The resulting estimates show that the estimated effect is always positive and significant and it becomes stronger for more restrictive sample definitions. This indicates that students at higher positions in the school-cohort ranking benefit more from being best in the class. This finding is consistent with psychological mechanisms like the impostor phenomenon. Students with the impostor phenomenon experience intense feelings that their achievements are undeserved and worry that they are likely to be exposed as a fraud (Clance, 1985). In my setup, students who are more likely to feel as an impostor are those students who are best in the class due to luck instead of true talent, i.e., those are relatively low positions in the school-cohort ranking.

The distance between the best and the second student in the class in terms of current performance may influence the effect of best in the class on future performance. On the one hand, best students who perform much better than the second student in their class may have higher self-esteem. On the other hand, best students who perform similarly to the second student in their class may exert more effort in response to competition or enjoy higher-quality peer effects from the second student in the class. I explore how the best in the class effect changes with the distance between the best and the second student in the class. To this, I perform alternative regressions in which I add the interaction of the best in the class dummy with a dummy for distance equal or higher than 0.221 (the median distance in the sample). I present the results of this exercise in Panel C of Table 10. Results show that the best in the class effect is higher for best in the class students who are closer to the second student in the class. This constitutes evidence in favor of the competition and better peer mechanisms.

6 Discussion

Being the best student in the class would be beneficial if it makes students have more self-confidence, exert effort in line with high expectations, or receive more attention and better treatment by teachers and peers. In contrast, being the best in the class would

Table 10: Heterogeneity

Panel A: Heterogeneity by Gender, Class Size, and School Quality. IV

	(1)	(2)	(3)
Best in class by male	0.08		
	(0.012)***		
Best in class by class size		-.010	
		(0.001)***	
Best in class by school quality			0.099
			(0.014)***
Best in class	0.084	0.313	0.143
	(0.013)***	(0.029)***	(0.011)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	247471	247471	247471
R^2	0.253	0.253	0.253

Panel B: Different Samples. IV

	up to 10th	up to 15th	up to 25th
	(1)	(2)	(3)
Best in class	0.166	0.144	0.132
	(0.01)***	(0.011)***	(0.011)***
Obs.	135020	205868	247471
R^2	0.173	0.224	0.253

Panel C: High Distance from Best to Second Student in the Class. IV

	(1)	(2)	(3)
Best in class by high distance	0.026	-.031	-.029
	(0.012)**	(0.011)***	(0.011)***
Best in class	0.118	0.123	0.125
	(0.016)***	(0.013)***	(0.013)***
Class fixed-effects	No	Yes	Yes
Test score dummies	No	No	Yes
Obs.	249309	249309	249309
R^2	0.245	0.253	0.253

Notes: Data is from INVALSI test for the years 2013-14 and 2014-15 together with test scores for years 2015-16 and 2016-17. The dependent variable is the mathematics test score in secondary school. All regressions include dummies for other positions in the class ranking, the individual characteristics described in tables 1 and 11, school-cohort fixed effects, and year indicators. Best in the class is instrumented using the theoretical probability of being the best in the class. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

be detrimental if it implies unbearable psychological pressure, if it makes students exert lower effort because they do not feel challenged or inspired by a better peer, if a better peer would have been helpful when studying or doing homework together, or if best in the class students are more likely to be bully victims. Hence, the question of whether being the best in the class is beneficial or detrimental does not have an obvious theoretical answer. Moreover, answering this question empirically is challenging because in most cases students are not assigned to classes randomly. Actually, in my data, I find evidence of positive assortative matching across classes within the same school. In this paper, I design a novel methodology to estimate the impact of being the best in the class on future performance. My methodology can be applied when experimental data is not available.

I exploit natural exogenous variation in class assignment within schools and I find that being the best in the class has enhancing effects on future performance. Results are robust to the study of different grades and to the use of reading rather than mathematics test scores. Interestingly, the effect of being second in the class is smaller, confirming that *best in the class* is different from other positions in the ranking. The estimated positive effect is stronger for males, students in smaller classes, and those in high-quality schools. It is also stronger for students in higher positions of the school-cohort ranking and for students who perform similarly to their best peer. My findings have implications in terms of the non-linearity of ranking effects at the extremes of the ability distribution and inform education policy: Excellent students may benefit from accommodation in a regular classroom where they are the best in the class rather than grouping with students of similar abilities.

References

ANGRIST, J. D., E. BATTISTIN, AND D. VURI (2017): "In a Small Moment: Class Size and Moral Hazard in the Italian Mezzogiorno," *American Economic Journal: Applied Economics*, 9(4), 216–49.

AZMAT, G., AND N. IRIBERRI (2010): "The importance of relative performance feedback information: Evidence from a natural experiment using high school students," *Journal of Public Economics*, 94(7-8), 435–452.

BERTONI, M., AND R. NISTICÒ (2018): "Rank Concerns, Peer Effects, and Ability Tracking in University. Evidence from a Randomized Experiment," .

BLACK, S. E., P. J. DEVEREUX, AND K. G. SALVANES (2013): "Under Pressure? The Effect of Peers on Outcomes of Young Adults," *Journal of Labor Economics*, 31(1), 119–153.

BOOIJ, A. S., E. LEUVEN, AND H. OOSTERBEEK (2017): "Ability Peer Effects in University: Evidence from a Randomized Experiment," *Review of Economic Studies*, 84(2), 547–578.

BROWN, G. D., J. GARDNER, A. J. OSWALD, AND J. QIAN (2008): "Does wage rank affect employees' well-being?," *Industrial Relations: A Journal of Economy and Society*, 47(3), 355–389.

BROWN, S., AND K. TAYLOR (2008): "Bullying, education and earnings: evidence from the National Child Development Study," *Economics of Education Review*, 27(4), 387–401.

CARD, D., A. MAS, E. MORETTI, AND E. SAEZ (2012): "Inequality at Work: The Effect of Peer Salaries on Job Satisfaction," *American Economic Review*, 102(6), 2981–3003.

CARRELL, S. E., R. L. FULLERTON, AND J. E. WEST (2009): "Does Your Cohort Matter? Measuring Peer Effects in College Achievement," *Journal of Labor Economics*, 27(3), 439–464.

CICALA, S., R. G. FRYER, AND J. L. SPENKUCH (2017): "Self-selection and comparative advantage in social interactions," *Journal of the European Economic Association*, 16(4), 983–1020.

CLANCE, P. R. (1985): *The impostor phenomenon: Overcoming the fear that haunts your success*, vol. 209. Peachtree Publishers Atlanta, GA.

COTTON, S. J., M. F. DOLLARD, AND J. DE JONGE (2002): "Stress and student job design: Satisfaction, well-being, and performance in university students," *International Journal of Stress Management*, 9(3), 147–162.

DENNING, J. T., R. J. MURPHY, AND F. WEINHARDT (2018): "Class Rank and Long-Run Outcomes," IZA Discussion Papers 11808, Institute for the Study of Labor (IZA).

ELSNER, B., AND I. E. ISPHORDING (2017): "A big fish in a small pond: Ability rank and human capital investment," *Journal of Labor Economics*, 35(3), 787–828.

ELSNER, B., I. E. ISPHORDING, AND U. ZÖLITZ (2018): "Achievement rank affects performance and major choices in college," *University of Zurich, Department of Economics, Working Paper*, (300).

FERKANY, M. (2008): "The educational importance of self-esteem," *Journal of Philosophy of Education*, 42(1), 119–132.

FIGLIO, D. N., AND M. E. LUCAS (2004): "Do high grading standards affect student performance?," *Journal of Public Economics*, 88(9-10), 1815–1834.

GENAKOS, C., AND M. PAGLIERO (2012): "Interim Rank, Risk Taking, and Performance in Dynamic Tournaments," *Journal of Political Economy*, 120(4), 782 – 813.

GILL, D., Z. KISSOVÁ, J. LEE, AND V. PROWSE (2019): "First-place loving and last-place loathing: How rank in the distribution of performance affects effort provision," *Management Science*, 65(2), 494–507.

GRIFFITH, A., AND K. RASK (2007): "The influence of the US News and World Report collegiate rankings on the matriculation decision of high-ability students: 1995–2004," *Economics of Education Review*, 26(2), 244–255.

GURYAN, J., K. KROFT, AND M. J. NOTOWIDIGDO (2009): "Peer effects in the workplace: Evidence from random groupings in professional golf tournaments," *American Economic Journal: Applied Economics*, 1(4), 34–68.

HORSTSCHRÄER, J. (2012): "University rankings in action? The importance of rankings and an excellence competition for university choice of high-ability students," *Economics of Education Review*, 31(6), 1162–1176.

KUZIEMKO, I., R. W. BUELL, T. REICH, AND M. I. NORTON (2014): "Last-place aversion: Evidence and redistributive implications," *The Quarterly Journal of Economics*, 129(1), 105–149.

MURPHY, R., AND F. WEINHARDT (2013): "The Importance of Rank Position. CEP Discussion Paper No. 1241.," *Centre for Economic Performance*.

——— (2020): "Top of the Class: The Importance of Ordinal Rank," *The Review of Economic Studies*.

SACERDOTE, B. (2001): "Peer effects with random assignment: Results for Dartmouth roommates," *The Quarterly journal of economics*, 116(2), 681–704.

TINCANI, M. (2017): "Heterogeneous peer effects and rank concerns: Theory and evidence," .

TRAN, A., AND R. ZECKHAUSER (2012): "Rank as an inherent incentive: Evidence from a field experiment," *Journal of Public Economics*, 96(9-10), 645–650.

VIDAL, J. B. I., AND M. NOSSOL (2011): "Tournaments Without Prizes: Evidence from Personnel Records," *Management Science*, 57(10), 1721–1736.

WHITMORE, D. (2005): "Resource and peer impacts on girls' academic achievement: Evidence from a randomized experiment," *American Economic Review*, 95(2), 199–203.

A Appendix

A.1 Institutional Background

The Italian education system is divided into primary school (grades 1 to 5), secondary school (grades 6 to 8) and high school (grades 9 to 13). Education is compulsory between the age of six (grade 1) and sixteen (grade 10). After secondary school, students start high school and follow one of three tracks (vocational school, technical school, lyceum). The school year starts mid-September and finishes mid-June. Education is compulsory from September of the year the student becomes 6 up to age 16 which implies that students who have not repeated any grade can drop out from school in grade 10 (second grade of high school). Students who repeat grades can drop out in lower grades as soon as they become 16.

The randomness of class assignment is at the heart of my estimation. In Italian primary schools, classes are formed in first grade when there is no comparable information on student performance (most students cannot read or write yet and in the vast majority of cases there are no interviews or psychological assessments) and the composition of classes is typically kept fixed up to fifth grade. New classes are formed in sixth grade when students move from primary to secondary school and their composition is kept fixed up to eighth grade. Again, principals of the new school typically do not have access to comparable performance information when they form classes. Even if they had, there are no official indications or directives on whether homogenous or heterogeneous classes should be formed. However, principals may use other observable characteristics (students' home address, parental education, previous school, number of siblings, special needs, etc.) to proxy for student future performance and use this information to form classes. Principals may also respond to parental requests to group their children with their friends'. Moreover, by the time that the ranking is measured students have already spent time in school, sharing teachers with their classmates. Hence, peer interaction, teacher traits, and school characteristics may generate a positive correlation of students' test scores within a class even if the class composition was random with respect to the pre-school ranking.

A.2 Other Student Characteristics

Table 11 describes daycare and kindergarten attendance, parental education and parental labor market status of students included in the regressions of being the best in the class in second grade on fifth grade performance, being the best in fifth grade on eighth grade, and being the best in eighth grade on tenth grade.

Table 11: Additional Descriptive Statistics

	Grades 2 to 5		Grades 5 to 8		Grades 8 to 10	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Attended daycare	0.393	0.488	0.336	0.472	0.274	0.446
Attended kindergarten	0.859	0.348	0.858	0.349	0.887	0.316
Mother primary school	0.015	0.12	0.017	0.131	0.012	0.109
Mother secondary school	0.214	0.410	0.243	0.429	0.228	0.419
Mother high school	0.437	0.496	0.43	0.495	0.446	0.497
Mother vocational school	0.069	0.253	0.078	0.268	0.086	0.28
Mother tertiary non-university	0.026	0.159	0.027	0.161	0.027	0.162
Mother university	0.24	0.427	0.205	0.404	0.201	0.401
Father primary school	0.018	0.135	0.021	0.142	0.015	0.120
Father secondary school	0.295	0.456	0.314	0.464	0.297	0.457
Father high school	0.401	0.49	0.389	0.487	0.397	0.489
Father vocational school	0.079	0.27	0.085	0.278	0.094	0.292
Father other tertiary	0.017	0.129	0.017	0.129	0.015	0.123
Father university	0.19	0.392	0.175	0.38	0.182	0.386
Mother unemployed	0.053	0.224	0.044	0.205	0.032	0.177
Mother homemaker	0.302	0.459	0.325	0.469	0.307	0.461
Mother white collar	0.445	0.497	0.429	0.495	0.453	0.498
Mother self-employed	0.084	0.278	0.086	0.281	0.094	0.291
Mother blue collar	0.114	0.318	0.114	0.318	0.113	0.317
Mother retired	0.001	0.032	0.001	0.034	0.001	0.038
Father unemployed	0.044	0.205	0.041	0.199	0.029	0.166
Father homemaker	0.004	0.059	0.003	0.056	0.003	0.058
Father white collar	0.444	0.497	0.441	0.497	0.455	0.498
Father self-employed	0.242	0.428	0.25	0.433	0.261	0.439
Father blue collar	0.261	0.439	0.256	0.437	0.241	0.427
Father retired	0.005	0.07	0.008	0.087	0.011	0.106
Year of the test	2013.504	0.5	2013.495	0.5	2014.523	0.499

Notes: This table presents averages and standard deviations (left and right column, respectively) for the samples used in the estimations.